

GIS-BASED RAPID POPULATION ASSESSMENT TOOL



Reference To Pending Prior Patent Applications

This patent application also claims benefit of pending prior U.S. Provisional Patent Application Serial No. 60/449,283, filed 02/21/03 by Wei Du for CENSUSVIEW: A GIS BASED SYSTEM FOR RAPID POPULATION ASSESSMENT IN CHES (Attorney's Docket No. HDM-4 PROV).

1 Overview (See Fig. 1)

The product contains nine categories of core functionality. Each of the following categories comprises the underlying structure of the rapid population assessment tool:

- Project Management
- Study Boundary Delineation
- Spatially Stratified Sampling
- Population Estimation
- GIS and Mapping Tools
- GPS Communication
- Analysis Tools
- GIS Interoperability and Web Service Integration
- Worldwide Population Portal Website

2 Project Management (See Fig. 2)

Projects are used to control sessions and manage project related documents when using the rapid population assessment tool. The project management engine contains functionality for users to manage project-related

information and data files. Outlined below are the required individual functions that make up the project management engine.

2.1 Create New Project

This function creates a new population estimation project. The following information is used to define a project:

- Project Name
- Project Description
- Creator
- Date Created
- Organization
- Data Projection
- Project Notes (arbitrary text entered by the user at any time)

*The Project Management engine can also incorporate other "metadata" information specific to a particular session.

2.2 Edit Existing Project

The application allows users to edit project information. This function gives users the ability to alter any of the seven function listed in section 2.1 for each saved project.

2.3 Delete Existing Project

The application allows users to delete an existing Project. The system will prompt a warning message to user before the project is removed from the system. If the targeted project is a currently opened project, the system will clear all files pertaining to the specific project.

2.4 Open Existing Project

The application allows users to open an existing Project. Individual projects should be opened after launching the application for use.

2.5 Save Project, Associated Settings, and Data Files

Projects are saved to the local file system on the users' computer. The application saves individual components that are Project specific. The ability to store the following is included:

- Project definitions: name, description, creator name, date last updated, project directory
- Boundary files (including exclusive areas) in shapefile, ASCII, or NMEA file formats
- Reference point files as shapefile, ASCII, or NMEA
- Sampling location files as shapefile, ASCII, or NMEA
- Survey location files as shapefile, ASCII, or NMEA
- Survey forms as HTML, Excel, or PDF file formats
- Sampling settings
- Survey data for individual study areas
- Population estimation results
- Last saved position and status. When the project is re-opened, its previous status will be restored to system.

2.6 View Project Summary Information

This function allows users to view all project information, including settings and associated data files.

2.7 Export Project Information

Each component of a specific project is exportable to file formats outlined in section 2.5.

2.8 Import Project Information

Each component of the project is importable from file formats outlined in section 2.5.

2.9 Print Project Information

This function allows users to print a Project's information in a well-designed report format.

2.10 General Undo / Redo

Each major function in the application is associated with an Undo/Redo function. The user is able to sequentially restore previous actions executed during the program. The application allows the user multiple levels of Undo/Redo for specific functions.

2.11 Re-Projection

Any spatial datasets used by the application have the ability to be re-projected for display and visualization of the data. The application allows users to display data in different accepted projections. This feature is necessary for accurate overlays of spatial data. The application contains standard coordinate systems for

users to choose from for re-projecting, taking into consideration commonly used formats throughout the world.

2.12 The User Map

The User Map is the general map display interface where spatial data is viewed by users and edited, where applicable. Individual requirements and functions relating to the User Map are outlined in further detail in section 6. When opened, each project will display its spatial layers and last extent according to the user.

2.13 Data Worksheet

The Data Worksheet is used for editing tabular data for individual projects. Features of the Data Worksheet include:

- Direct data entry (and also includes “intelligent” features such as combo/drop-down boxes, checkboxes, pre-selected data).
- Cut/copy/paste functions using standard Ctrl-X, Ctrl-C, and Ctrl-V keyboard shortcuts.
- Change data view, by moving columns. Also includes hiding columns and similar features.
- Ability to sort data by fields and/or rows
- Reorder data. The application allows users to move a rows up or down from the worksheet view.

3 Boundary Delineation (See Fig. 3)

The boundary delineation functionality provides comprehensive tools for users to define study area boundaries, exclusive areas, and reference points. The product allows these spatial components to be

downloaded from GPS devices, entered manually using spatial coordinates, or be digitized on screen by user. Information pertaining specifically to GPS communication functions is described in Section 8.

3.1 Load Boundary Data from GPS

The application is able to load data directly from compatible GPS units. Functions include:

- Loading GPS data directly from GPS unit as full Study Boundary Area or as an Excluded Area
- Loading GPS data directly from GPS unit as a subset of a full Study Area or Excluded Area.
Subsets of full boundaries must be indicated in some way (preferably color-coded). The application is also be able to connect multiple parts to create a single Study Boundary or Excluded Area.
- The boundary data downloaded from GPS units can be displayed in the User Map.

3.2 Load Boundary Data from GPS/GIS Files

For data gathered using GPS units that are not compatible with direct connect functionality, data files are supported. The application is able to incorporate and use GPS data files for boundary creation and display, meeting the requirements listed in Section 3.1.

3.3 Edit Boundary Data on Worksheet

Vertices of all boundary polygons are editable using the application's data worksheet. Functionality specific to boundaries should include the ability to load boundary data directly from a GPS device or associated GPS format file.

3.4 Key-in Boundary Points into Worksheet

For users without a GPS device or without a GPS connection, the application provides the flexibility to enter boundary data via the keyboard for display in the User Map.

3.5 On-Screen Digitizing of Study Area Boundaries

The application allows users to create and edit boundaries using on-screen digitizing methods. This function also allows users to load satellite imagery for viewing in the User Map to perform the onscreen assessments. Functions provided include:

- The ability to load georeferenced satellite imagery if available to assist with digitizing tasks.
- Create boundaries and exclusive areas by on-screen digitizing.

3.6 Create Study Area, Sub-Areas, and Excluded Areas

Study Areas, Sub Areas, and Excluded Areas are able to be created using the boundary creation methods described above. We define the term "SubArea" to mean areas within a delineated Study Area from a single project. The application accommodates multiple Sub Areas and Excluded Areas. Excluded Areas are regions within a given Study Area designated by the user that are removed for sampling purposes, such as areas representing lakes within a given Study Area.

3.7 Export Boundary Data File

The application allows users to export boundaries into different file formats. Formats include the following:

- Shapefile
- ASCII
- NMEA
- GML

3.8 Study Area Analysis

The application allows users to compute the total area of Study Area polygons. The application may perform other GIS-related analysis on study areas as needed or requested. Analysis results may be saved as attributes for the particular area.

3.9 Load / Edit / Save Reference Points

Reference points allow users to create or show spatially important areas as points in the User Map. These points incorporate GIS functionality similar to Study Areas, where applicable. This includes the ability to name specific reference points and store related attribute information.

3.10 Spatial Editing Functionality

Study Areas, Sub-Areas, reference points, and other spatial objects are editable on-screen using the User Map and associated GIS functions. Specific functions include:

3.10.1 Select spatial objects for editing

This function allows users to select a spatial object for editing. The spatial object includes boundary polygons, exclusive boundaries, polygon vertices, reference points, boundary buffers, sub sampling areas, and other spatial objects. The spatial objects on map can be selected in two ways: (a) by a single map click, and (b) by a user-defined polygon. Where possible, the user should be able to select more than one object at a time.

3.10.2 Add / Delete vertices of polygon objects

This function allows users to add or delete vertices of a polygon object.

3.10.3 Move vertices of polygon objects

This function allows users to move a selected vertex of a polygon object

3.10.4 Create a shortcut between two selected vertices

This function allows users to simplify an arc by creating a straight line between two selected vertices of an existing polygon object.

3.10.5 Snap selected end nodes together

This function allows user to select two nodes (end points on any arc) and combine them into a single node.

This node should correspond to one of the selected nodes (and should be consistently applied). The end

nodes come from the same arc, which will create a single polygon. The end nodes may come from different line objects, which will create a single line.

3.10.6 Delete a selected spatial object

This function allows user to delete any selected spatial object (points, lines, polygons, and vertices) within the User Map.

3.10.7 Move a selected spatial object

Spatial objects are movable by changing their position within the User Map. For objects other than points, moving should not change original shape.

4 Spatially Stratified Sampling (See Fig. 4)

The spatially stratified sampling functionality provides the foundation for area and point-based population estimation. A unique spatially stratified sampling methodology has been created.

Spatially stratified sampling methods can be classified as:

- **Area-Based Sampling**
 - Systematic
 - Random
- **Point-Based Sampling**
 - Systematic
 - Random

Area-based sampling methodology is used by the Quadrat Population Estimation technique. Point-based sampling method is utilized by the T-Square and Point Quarter Population Estimation techniques.

4.1 Sampling Settings

The software allows users to create comprehensive sampling schemes. Each Study Area has a single sampling scheme associated with it. In a project with more than one Study Area (sub-areas), each may have a distinct sampling scheme.

The procedure for sampling setting includes the following steps:

- 1) Select one or more Study Areas for sampling
- 2) Select a sampling method: Quadrat, T-Square, or Point-Quarter methods (others possible)
- 3) Define sampling size. For point-based estimation methods this is simply the number of points to be generated. For area-based estimations, the width and length of each sampling area must be defined. There may be additional methods required in the future (such as transect-based methodologies).
- 4) Select a sampling type: systematic or random

4.2 Define Boundary Buffer

In order to ensure usability of all sampling locations, the system allows users to create a buffer in which no sampling locations will be located. If a buffer is created, all sampling operations will be restricted to areas in the polygon not covered by the buffer.

- For Study Area boundaries, the buffer shall be inside the chosen polygon.
- For excluded areas, the buffer shall be located outside the area, surrounding it.

4.3 Define Sub Sampling Areas for Multiple Frequency Sampling

This function allows users to define sub-sampling areas by user-defined polygons within a given Study Area. Currently there is a three sub-area limit in the prototype; however, the number of sub-areas may be unlimited.

4.4 Sampling Method Setting and Parameter Management

Meaningful defaults are given for each Sampling Method, where possible.

For advanced users, the system also provides tools for comprehensive sampling method setting, including the ability to alter detailed sampling algorithms (such computation iteration), where applicable.

4.5 Define the Fields for Data Collection

The system allows users to define fields (column names within the data worksheet) for data collection. To simplify this process, two types of fields are defined in the system: 1) a NUMBER type field and 2) a TEXT type field. The system also provides some basic fields for use by default.

4.6 Execute the Spatial Sampling

The system generates sampling locations for the user. Their location depends on the sampling parameters chosen. Sampling locations and identifying labels are displayed as a layer on the User Map.

4.7 Sampling Result Management

The system is able to export results (sample locations) into any of the following formats:

- Shapefile
- ACSII
- NMEA
- GML

The system also is able to upload sampling outputs to a connected GPS unit. The system divides the total sampling scheme results into parts for upload to multiple GPS units. This allows for collaborative division of survey points among multi-member teams.

4.8 Create the Survey Form

The system is able to create a field survey form for data entry in HTML, Excel, or PDF formats after the sampling scheme is executed. The survey form provides sampling location information, including directions from reference points to each sample location, and blank tables for data recording. The system also allows users the ability to print the generated forms for use out in the field.

The system is able to select sub-sets for the creation of survey forms. This enables a project coordinator to send multiple workers into the field to collect data for specific locations.

4.9 Sampling Scheme Management

The application is able to manage each Sampling Scheme created by the user. A Sampling Scheme includes the following information:

- Sampling-specific parameters (number of points, confidence interval, etc.)

- Survey fields to be collected
- Estimation model used

Management of Sampling Schemes include the following functions:

- Save to file under a name prescribed by the user.
- Load from file with ability to browse associated directories.

The system may include pre-defined Sampling Schemes for the user to load. These pre-set schemes may be chosen by the user to simplify procedures.

5 Population Estimation (See Fig. 5)

5.1 Input Field Survey Data for Population Estimation

Data collection at each of the generated Sampling Locations is necessary for population estimation. The collected data must be input to the program in one of the following ways:

- Input by tabular worksheet. This function supports conversion from paper surveys.
- Individual data entries through an input format. Here, the application allows users to navigate the data using “next” and “back” functions.
- Import of structured text (CSV or potentially GML) file.
- Import of GIS data file (Shapefile).
- On-screen input using User Map (see examples below).

All data input to the application is editable using a Data Worksheet (2.13).

5.2 Select Data Fields for Population Estimation

Where the Project contains survey data in multiple fields, the user is able to select a limited number of fields to use for each estimation calculation.

5.3 On-Screen Data Collection Functionality Examples

The application provides effective procedures for users to perform tasks at different sampling locations, navigate through survey points, generate distance measurements, and store them in associated data tables automatically.

5.3.1 On-Screen Survey for T-Square Method (See Fig. 6)

This function automatically zooms to a selected sampling point and highlight it. The system allows users to collect two map points (such as identifiable dwellings) for the distance computation. After the first point is chosen, the system automatically draws reference lines to help user to determine the second map point with an appropriate angle.

5.3.2 On-Screen Survey for Point Quarter Method (See Fig. 7)

This function automatically zooms to a selected sampling point and highlight it. Two reference lines are drawn at the sampling point to represent four quads for map point determination. The system allows users to collect four map points for distance computation. Based on these reference lines, users can collect one map point (the nearest housing unit) from each quad. The system automatically computes the distance between the map points and the sampling point, and store results in the data table.

5.3.3 On-Screen Survey for Quadrat Method (See Fig. 8)

This function automatically zooms to currently selected sampling quad and highlight it. Based on satellite imagery, user will determine how many housing units are within the sampling quad area, entering this number into the system.

5.3.4 Enter Average Number of People Per Housing Unit (Optional)

When actual population data for the targeted housing units is not available, users can input the average number of people per housing unit from historic data or other sources to perform the population computation. This function allows users to complete rapid population estimation for a given area when minimal data is available.

5.4 Execute the Population Computation

Based on user-defined data or data collected from a Study Area, the population computation can be executed. The system will display the sampling scheme summary and population results in a report format on screen. The report is also viewable in HTML and/or PDF formats.

6 User Map and GIS Functionality (See Fig. 9)

The application provides the basic GIS-related tools needed to map and analyze the spatial data used by the application.

6.1 Base Data

The application includes a number of base data layers for users to display. Some data layers will be automatically added to a User Map when an existing project is opened or new one is created. Currently, a data layer depicting countries of the world is the only data source included. The ability to include additional standard data layers is also included.

6.2 Layer Management

Layer management functionality includes:

Add a Layer, including both vector and raster-based datasets. Supported vector GIS formats include the following:

- ESRI Shapefile
- OGC WFS format (GML)

Layer visibility and scale control: Methods exist within the application to allow for turning on and off the visibility of specified data layers. Also, the scale at which specific data layers appear within the User Map should be controllable. Functions shall also exist to remove one or more layers from the User Map view.

6.3 Map Navigation

General navigation around the User Map includes these common GIS functions:

- Zoom In by map click and drag box
- Zoom Out by map click (and a drag box)
- Gazetteer: Zoom by place name, depending on place names/data layers available. The Gazetteer can be used even if the vector Layer is not actually on the current Map.
- Zoom to a defined project object (such as boundary extent)

- Pan / Re-center
- Move back to a previously defined map position (at least 20 map positions should be stored and re-cycled)
- Move forward to a previously defined map position

6.4 Thematic Mapping

The system allows users to define the following parameters for vector-based data layers:

- Custom attribute and color ramp setting
- Custom map style and pattern setting
- Labeling by attributes
- On-the-fly custom points (graphics) to be drawn on the on User Map
- Map legend displaying all current data layers and/or symbology

6.5 Contour Map Generation

This function is designed to create a contour map based on point data (Z values). This function can be used to create population density maps based on point data collected from a study area.

6.6 Print and Export Map

Functionality includes:

- Print Map at any scale
- Export map to different raster formats: png, jpg, and gif
- Export individual vector data layers to ESRI shapefiles, wherever applicable.

6.7 Save Map Setting and View

This function allows user to save current map setting (data layers used and their individual settings, including current extent and position on the User Map).

6.8 Load Existing Map

The user is able to load an existing map including its associated data layers to the User Map.

6.9 Info Tools

Other GIS-based tools implemented in the User Map include, but are not limited to:

- Info Tool: display attribute information for selected spatial objects
- Index Map: display current map extent on a world map
- Map Reader (must attempt Section 508 compatibility at the mapping level)

7 GPS Communication (Fig. 10)

GPS communication functionality encompasses two aspects:

- **Direct Communication:** For GPS devices with open specifications for application protocol, the product will communicate directly with the device for the upload and download of waypoint and/or route (line, polygon) data. At this time, manufacturers supporting this method include Garmin, Magellan, and Lowrance. More may be added as identified.

- **File-Based Communication:** For GPS devices without open specification for application protocol, the product will rely on a common file interface to communicate with the devices. File interface will depend on the unit / manufacturer, but will resemble the NMEA standard format. GPS manufacturers known to operate using this method include Trimble.

8 Analytical Tools (Fig. 11)

When the population estimation is performed, many comprehensive analyses can be applied. Analyses can currently be grouped into five major areas. Listed below are the application's general requirements. Each analysis is available in three forms:

- **Tabular:** This functions as the compiled Data Worksheet, but in a read-only mode.
- **Graphical:** Charts and/or graphs are generated.
- **Report:** Analyses are printable and may include information interpreting the results. Reports are created in HTML or PDF formatted text.

8.1 Population Profile Builder

This function allows users to build epidemiological, demographic, and social & economic profiles for resulting population. Users are able to define a profile type, profile name, and associated ratio or rate for each population interval. The ability to add/remove/edit profiles from the system is also incorporated.

8.2 Resource Allocation Analysis

This function allows users to analyze itemized resource needs for relief activities by defining quantities and costs of relief supplies and apply them to a specific Study Area.

8.3 Epidemiological Analysis

This function allows users to analyze epidemiological impacts for the displaced population, and predict outcomes.

8.4 Demographic Analysis

This function allows users to analyze demographic patterns for the population within a Study Area and predict changes of demographic composition.

8.5 Social and Economic Analysis

This function allows users to analyze social and economic impacts for the specific population and predict changes in these attributes.

9 GIS Interoperability and Web Service Integration (See Fig. 12)

The application is able to utilize GIS data for the User Map through various web services specifications, including the required parts of the WMS (web mapping service) client specification. No data serving is required by this application, except to upload data and results to a future web page design.

The application is designed with connectivity to and from remote sources in mind.

10 Worldwide Population Web Portal (See Fig. 12)

The goal of the worldwide population web portal is to promote world communities to exchange and share population information, supporting efforts for humanitarian relief activities. In general, the web portal includes:

- Displays worldwide events for population studies
- Refugee and humanitarian relief news
- Platform for population data exchange
- Resources and tools for humanitarian emergency planning and responses
- Web hosting infrastructure and capabilities

11

Technologic Summary of Methodologies in CensusView™

CensusView™ is a powerful rapid population assessment system. The system uniquely integrates geographic information system (GIS) and global positioning system (GPS), remotely sensed imagery (including satellite imagery), advanced population estimation

research, and spatially stratified sampling methodologies for rapid population assessment. The system can significantly increase productivity and accuracy of field population assessment, reduce field population survey risk during complex humanitarian emergencies (CHEs), and help in making effective decisions on relief efforts. The system is the first tool that standardizes population estimation procedures crossing international boundaries, and that integrates world-wide geospatial information with population estimation through interoperable web services crossing multiple geospatial data platforms, international boundaries, and world-wide relief organizations.

Technical Highlights:

- Advanced GIS object model and multiple GIS package integration capabilities
- Most comprehensive population estimation methods based on real world experience
- Multiple spatial sampling models, covering area-based and point-based sampling methodologies
- Multiple-area samplings with different sampling frequencies
- Reduction of unusable sampling points or quads near specified population area boundaries
- Multiple population estimation methodology: quadrat, t-square, and point quarter methods
- Effective PDA and Desktop Software Communication and integration

CensusView™ can assist CHE specialists with the following functions rapidly and effectively:

- Investigate disease outbreaks and reduce infectious disease transmission

- Identify vulnerable populations and tailor programs to meet population specific needs
- Calculate rates of disease in surveillance systems to direct programs
- Assess the availability of water, fuel, and other resources
- Assess the population's accessibility to health, sanitation, and other services

See Fig. 13

A. Spatially Stratified Random Sampling Algorithm (See Fig. 14)

1. The novel spatially-random stratified-sampling methodology has the ability to simulate complex boundary conditions in the field, including
 - a. Develops boundary "inner buffer polygons" to avoid creating unusable sampling points
 - b. Generates "exclusive area polygons", to define unwanted areas for sampling within study area polygons
 - c. Generates "cluster boundary polygons" for different frequency samplings, to enable complex population cluster sampling.

The methodology ensures that the systems' random sampling has the following characteristics:

- the random sampling locations are spatially stratified, ensuring that the sampling location geographically represents the study areas, and

- the spatial bias of randomness is minimized, ensuring that the designated area within the study area boundary has equal probability for random sampling

The spatially-stratified random-sampling methodology is implemented by the following procedure:

1. Create a systematic grid, containing exact number of target grid points within the study area. An example grid layout is shown in Fig. 15.
2. Systematically divide each grid cell into an infinite number of very fine sub-grid cells (for example, $1,000 \times 1,000 = 1,000,000$ cells). For each grid cell, sub-grid cells are created from the left to right and from the bottom to top. Higher numbers of fine grid cells helps to ensure that any given cell within the study area has an equal possibility for receiving a random sampling.
3. Assess and list all sub-grids that *intersect* the study area. The list of intersected cells has the same sequence as when the grid and sub-grid cells were created.
4. Divide the sub-grid list into a number of sections. Each section contains equal numbers or near equal numbers of sub-grid cells. The total number of sections is equal to the exact number of targeted sampling locations.
5. Randomly select a sub-grid cell from each section (group) of sub-grid cells defined above.
6. This method is computationally intensive. An optimal computation procedure is developed in Java to achieve the computation goal with very good performance.

B. The Most Comprehensive Population Estimation Methodology Integration (See Fig. 16)

Traditionally, population estimation under complex humanitarian emergencies is carried out by using a manual procedure from data collection to population computation. There is no such tool that is capable of standardizing the procedure, then managing and integrating information and technologies from different steps of the population estimation process together. In most cases, data are entered into spreadsheet type software, and population computation is done in an inconsistent fashion. CensusView™ is the first system that integrates and computerizes all of the following population estimation methodologies and procedures. Population estimation methods implemented into current version of CensusView™ include

- Quadrat Method
- T-Square Method
- Point Quarter Method
- Transect Method

Geographic information systems (GIS), global positioning systems (GPS), and satellite imagery technologies are integrated into different steps of the methods used by CensusView™, making the entire process easy to learn, operate, manage, implement, and analyze results in the field.

C. Satellite Imagery Based On-Screen Population Survey Methodology

When deploying personnel into critical areas is not feasible, CensusView™ can load commercially available satellite imagery into its GIS framework, allowing users to demarcate dwellings and apply spatially-stratified sampling procedures with predefined estimates to ascertain the overall population in a defined area. Results from this "On-Screen" Survey are achievable without the need for a field survey. The time consumption for the population assessment can be reduced from a number of days or weeks to a few minutes. The system can help users automatically navigate through sampling locations for distance measurement automation, data input, and visualization. The system creates "roof-top accuracy" for associated distance measurements. During the assessment process, the system draws reference lines to guide the user to pick housing units, buildings or other structures based on the user-selected population method.

For example, in the Point Quarter method, a user clicks on the nearest house in each quad (four clicks on the map). The system automatically computes the distances among the sampling point and survey points, and adds this information into the database. The option of "map zoom" allows user to automatically zoom to selected sampling location.

See Fig. 17

For the T-Square method, only two survey points need to be defined, and other procedures follow the same methodology as the Point-Quarter method described above. The system draws a reference line after the first point is defined. Users use the reference lines to make sure that the angle among the sampling point, the first nearest house to the sampling point, and the first nearest house to the first survey point is greater than 90 degree.

See Fig. 18

CensusView is the only system that integrates measurement procedures, distance computations, sampling rules, and methods with high-resolution satellite imagery technology for population estimation.

Further details of CensusView are disclosed in Fig. 19.